

Escudero, Tomás (2016). La investigación evaluativa en el Siglo XXI: Un instrumento para el desarrollo educativo y social cada vez más relevante. *RELIEVE*, 22 (1), art. 4. DOI: <http://dx.doi.org/10.7203/relieve.22.1.8164>

Revista ELección de Investigación
y EValuación Educativa



e-Journal of Educational Research,
Assessment and Evaluation

La investigación evaluativa en el Siglo XXI: Un instrumento para el desarrollo educativo y social cada vez más relevante

Evaluation research in the 21st century: a more and more relevant tool for the educational and social development

Escudero, Tomás

Universidad de Zaragoza

Resumen

Tras una amplia revisión de las recientes publicaciones alrededor del tema, se analiza y valora la actual situación de la investigación evaluativa, como instrumento estratégico para la toma de decisiones de desarrollo y mejora de la sociedad y de la calidad de vida de los ciudadanos, en ámbitos diversos como la educación, la sanidad, la economía, la cultura, la protección social, las políticas públicas, etc. Se describe y fundamenta la identidad científica de la investigación evaluativa actual, incidiendo en su carácter transdisciplinar, en el auge de la evaluación de organizaciones e instituciones, en su apoyo en metodologías diversas y en la importancia de las estrategias participativas. Se destaca también la utilidad y el uso apropiado de las evaluaciones como objetivo prioritario de este tipo de investigación, apoyándose siempre en principios y normas éticas y de calidad científica y los correspondientes estudios metaevaluativos.

Fecha de recepción
21 Diciembre de 2015

Fecha de aprobación
22 de Marzo de 2016

Fecha de publicación
23 de Marzo de 2016

Palabras clave:

Investigación evaluativa, Desarrollo social, Disciplina transversal, Metodologías diversas, Estrategias participativas, Utilidad y uso de la evaluación, Normas ético-científicas, Metaevaluación

Abstract

After a wide revision of the recent publications about the subject, the present situation of evaluation research is analysed as a strategic tool for the decision making of the development and improvement of society and citizens' quality of life in varied sectors such as education, health, economy, culture, social protection, public policies, etc. The scientific identity of the evaluation research is described and founded upon, stressing its transdisciplinary character, the rise of the evaluation of organizations and institutions, and the use of different methodologies and the importance of participative strategies. Also outlined is the utility and appropriate use of the evaluations as priority object of this kind of research, relying always on principles and ethical rules of scientific quality and on the corresponding meta-evaluative studies.

Reception Date
2015 December 21

Approval Date
2016 March 22

Publication Date:
2016 March 23

Keywords:

Evaluation research, Social development, Transdisciplinary discipline, Different methodologies, Participative strategies, Utility and use of the evaluation, Ethical-scientific rules, Meta-evaluation.

La investigación evaluativa se ha convertido en las últimas décadas en un enfoque muy importante de actividad investigadora en todos los ámbitos científicos y de manera especial en ámbitos relacionados

con los programas y políticas sociales, en general, y en educación en particular. Cada vez son más los investigadores que trabajan en este campo e incorporan en sus

Autor de contacto / Corresponding author

Escudero, Tomás. Universidad de Zaragoza. Departamento de Educación, Instituto de Ciencias de la Educación. C/Pedro Cerbuna, 12. 50009 Zaragoza (España). tescuder@unizar.es

planteamientos sus principios, criterios y estrategias metodológicas.

La investigación evaluativa, también conocida en la literatura especializada como evaluación, se mantiene en constante evolución, pero tuvo su periodo central de desarrollo en la segunda mitad del siglo XX, en buena medida, como respuesta obligada a la necesidad de analizar, valorar y mejorar las políticas, acciones, planes, programas, instituciones y sistemas en las que se apoyan las sociedades modernas para desarrollarse y mejorar la calidad de vida de sus ciudadanos (Escudero, 2011).

Durante el aludido periodo, creció el impulso, desde los gobiernos y otras organizaciones, de planes y programas ambiciosos de cambio y mejora social, relacionados con la educación, la salud, la cultura, la protección social, la propia investigación científica, etc., que requerían ser analizados y valorados en profundidad, para conocer bien su funcionamiento, su eficacia, su eficiencia, y para poder tomar decisiones sobre su continuidad y su transformación. Por ello, la investigación evaluativa se establece como un *mecanismo estratégico* de cambio social justo y equitativo (Cook, 2015; Schawndt, 2002) y una aliada obligada de los responsables políticos y sociales para optimizar sus acciones y sus decisiones. Un claro ejemplo de esta orientación hacia la mejora social lo encontramos en el número de Verano 2015 de la importante revista “*New Directions for Evaluation*”, dedicado monográficamente a evaluaciones de programas para mejorar la justicia social y resolver problemas de grupos desfavorecidos. Por ello, resulta lógico el planteamiento de Thomas y Madison (2010) proponiendo y defendiendo la integración de la disciplina *justicia social* en los programas de formación de evaluadores. Programas que, tal como indican La Velle y Donaldson (2010), han tenido un gran crecimiento en los últimos años. En este contexto, es lógico, tal como se hace hoy (McClintock, 2003; Calderon, 2004), definir al evaluador como un *investigador agente del cambio*.

En línea con esta realidad nos encontramos con la declaración de 2015 como *año internacional de la evaluación*, en la III Conferencia Internacional sobre Políticas Nacionales de Evaluación, que tuvo lugar en Sao Paulo entre el 29 de septiembre y el 2 de octubre de 2013. La propuesta fue de la asociación “EvalPartners”, un movimiento internacional global de evaluadores para fortalecer las capacidades nacionales en este terreno, con el objetivo declarado de utilizar la evaluación para mejorar la vida de la gente, incidiendo en la mejora de las políticas públicas.

Esta gran incidencia de los programas sociales en el desarrollo de la investigación evaluativa, ha hecho que con frecuencia se siga hablando de *evaluación de programas* como denominación de este tipo de investigación (Smith, 2010). Sin embargo, el hecho de que el objeto de las evaluaciones se haya ampliado más allá de los programas, por ejemplo, políticas en general, instituciones, sistemas, organizaciones, recursos, personal, etc., ha servido para acuñar y consolidar la denominación más global y ajustada de *investigación evaluativa*.

Esta investigación adquiere una gran relevancia y desarrollo en ámbitos como la educación, la salud y la intervención y el bienestar social, por la gran importancia que tienen para el funcionamiento y desarrollo de las sociedades modernas, pero debemos puntualizar que el alcance de este tipo de investigación es transversal e incide igualmente en el resto de disciplinas y ámbitos de relevancia social, como economía, cultura, ciencia, investigación, deporte, información, atención ciudadana, etc. Así, por ejemplo, aprovechando la puesta en marcha de la Revista de Evaluación de Programas y Políticas Públicas, Muñoz y col.(2013), acompañando a una revisión del tema en el ámbito económico, defienden la creciente necesidad de este tipo de investigación en materia de políticas públicas en España y en el conjunto de la Unión Europea. Por estos motivos, toma cada vez más relevancia el conocido planteamiento de Scriven (2003)

definiendo y defendiendo a la investigación evaluativa como disciplina transversal o transdisciplina, pues su ámbito de análisis y mejora se extiende no sólo a objetos muy diversos de evaluación, sino también a muchas disciplinas.

Identidad y fundamentación científica de la investigación evaluativa actual.

Cuando hablamos de investigación evaluativa estamos tratando con un concepto complejo, con distintas orientaciones, que ha seguido y sigue un proceso sistemático de evolución. En el ámbito educativo, por ejemplo, es muy interesante observar este proceso de desarrollo desde cuando la evaluación se identificaba con la medición de características personales, hasta la actual investigación evaluativa (Escudero, 2003).

Por este motivo, pretender definir la investigación evaluativa con una simple frase, delimitándola de manera precisa de otros enfoques o ámbitos investigadores resulta difícil, porque existen muchas zonas de contacto, que hacen borrosa la distinción. Solamente resaltando los elementos más específicos de la investigación evaluativa seremos capaces de establecer un *perfil definitorio válido y útil* para encuadrar el trabajo científico en este campo de desarrollo creciente.

Lo primero que se debe resaltar, pues es un elemento definitorio primordial, es que la investigación evaluativa se encuadra en un *contexto de cambio* y más concretamente, en un *contexto de cambio social*. Se trata, de un *enfoque reactivo* de la investigación, dentro de la *corriente de pensamiento paradigmático de base crítica*, que busca ofrecer soluciones a problemas concretos, desde una posición *pragmática y contextualista*, desde la que se defiende una noción situacional del desarrollo social, en vez de unos métodos estándar fijos como base y apoyo al desarrollo social emancipatorio (Bredo, 2006; Escudero, 2009).

Este enfoque paradigmático, orientado hacia la resolución práctica de problemas, tiene al igual que el hermenéutico su apoyo en el construccionismo social, pero usa también métodos cuantitativos como el experimental,

así como indicadores de funcionamiento, que apoyen el diseño de intervenciones eficaces. Siguiendo los principios del *pragmatismo* y el *contextualismo*, se buscan programas y proyectos que «funcionen bien» dentro de un contexto real particular, no siendo el propósito fundamental la descripción y construcción teóricas, sino la toma de decisiones para la gestión y la planificación, la solución de problemas y la construcción práctica de programas, dentro de la tradición investigadora aplicada (Escudero, 2011). Esta realidad hace que en la investigación evaluativa, la indagación completa de procesos y resultados y de sus significados, obliga a la utilización de planteamientos diversos y mixtos (Sondergeld & Koskey, 2011).

Vemos que la investigación evaluativa se plantea siempre en un contexto real, al servicio de las políticas sociales, en definitiva, al servicio del cambio y del desarrollo social y que, por lo tanto, es un instrumento intermedio, de ayuda a otras disciplinas y ámbitos como la educación, la sanidad, la cultura, el bienestar social, etc., pues busca directamente el objetivo de ofrecer las mejores opciones de acción entre todas las alternativas posibles. Se trata de ayudar a resolver problemas que aparecen en estos ámbitos. La investigación evaluativa, en suma, se mueve en un contexto de resolución de problemas. Y en este sentido, el evaluador se ve obligado a utilizar el conocimiento teórico en la resolución de problemas prácticos, jugando el importantísimo papel de facilitador de la integración de la investigación básica y la práctica (Urban & Trochim, 2009).

Otro elemento identificativo importante de la investigación evaluativa es que aborda el análisis de todo tipo de información, tanto cualitativa como cuantitativa, en muchos casos con grandes volúmenes de datos, desde perspectivas diversas, apoyándose en metodologías de investigación múltiples y flexibles (Escudero, 2009; Maxcy, 2003).

En este contexto también hay que destacar, por lo que significa de expansión de la evaluación a una nueva perspectiva, *la evaluación para el desarrollo o evaluación*

capacitadora (empowerment evaluation), impulsada por Fetterman como un enfoque, de base democrática, de participación de los implicados en el programa evaluado, para promover la autonomía de los mismos en la resolución de sus propios problemas (Fetterman, 2001b; Fetterman, Kaftarian & Wandersman, 2015).

Actualmente, también tenemos una corriente bastante activa de evaluadores que se adscriben a la visión denominada *evaluación basada en la teoría*, y que defienden la necesidad del evaluador de contar previamente, y analizar críticamente, con el modelo lógico (Renger & Hurley, 2006; Wasserman, 2010) o teórico que fundamenta el programa, proyecto, las políticas y las intervenciones a evaluar (Christie, 2003; Donaldson y Gooler, 2003; Donaldson, 2007; Donaldson & Lipsey, 2006; Donaldson & Scriven, 2003; Mark, 2003; Weiss, 2004).

Esta visión recoge planteamientos diversos, pero defiende la contextualización y fundamentación específica de cada investigación, de acuerdo con las características particulares del problema, programa u objeto a evaluar, de los implicados en el mismo y de las posibles vías de acción más eficaces. La adaptabilidad, la flexibilidad, la integración de enfoques y el eclecticismo metodológico son elementos definitorios de esta aproximación de la investigación evaluativa, que parte del desarrollo de una teoría fundamentadora del programa o de la intervención, desde distintas fuentes y procedimientos, desde donde plantearse las cuestiones de evaluación prioritarias a responder.

Otro enfoque vigente en los últimos años es el de fundamentar las evaluaciones desde la perspectiva teórica de la complejidad de los objetos y programas evaluados, tanto para definir los objetivos como las aproximaciones metodológicas. Walton (2014) nos ofrece una revisión de la fundamentación teórica de este enfoque, de sus características principales y de las consecuencias de su aplicación en la última década, que confluyen en buena medida con otras aproximaciones que resaltamos en

este trabajo, destacando aspectos como el uso de métodos múltiples y mixtos y el de las metodologías participativas.

Urban, Hargraves y Trochim (2014) defienden la investigación evaluativa como un proceso evolutivo en distintas fases, en correspondencia con el propio proceso evolutivo de los programas evaluados, de manera que en cada fase se plantean objetivos y metodologías diversas. En concreto, nos hablan de cuatro fases: una primera de *inicio* del programa, donde se evalúa la *respuesta al proceso* por parte de los participantes; una segunda de *desarrollo* del programa, donde se evalúan los *cambios* que produce en los individuos, contexto, etc.; una tercera donde se analiza la *estabilidad* del programa, *controlando* y evaluando su eficacia en *comparación* entre grupos; y finalmente una cuarta fase de *diseminación* del programa, en la que se evalúa la capacidad de *generalización* del mismo.

Una valoración comprensiva de los planteamientos analizados, nos conduce a una posible definición de la actual investigación evaluativa, como “un tipo de investigación aplicada, que incide sobre objetos sociales, sistemas, planes, programas, participantes, instituciones, agentes, recursos, etc, que analiza y juzga su calidad estática y dinámica según criterios y estándares científicos rigurosos múltiples, externos e internos, con la obligación de sugerir acciones alternativas sobre los mismos para diferentes propósitos como planificación, mejora, certificación, acreditación, fiscalización, diagnóstico, reforma, penalización, incentivación, etc.” (Escudero, 2006, p. 271).

Consolidación como transdisciplina

En la *perspectiva transdisciplinar*, la investigación evaluativa, manteniendo su idiosincrasia como disciplina autónoma, proporciona herramientas y apoyos básicos a otras disciplinas muy diversas (Scriven, 2003). Se puede hablar, por tanto, de dos componentes: el conjunto de campos de aplicación de la evaluación y el contenido de la propia disciplina. Algo parecido a lo que

ocurre a disciplinas como la estadística y la medición. En definitiva, en la práctica, la investigación evaluativa es una disciplina que incluye sus propios contenidos y los de otras muchas disciplinas que son objeto de estudio. Y como señala Jacob (2008), las evaluaciones requieren la colaboración estrecha entre especialistas en distintos ámbitos y dominios del conocimiento.

Fortaleciendo este planteamiento, investigadores relevantes como Stufflebean (2001a), han defendido e impulsado programas de formación académica en investigación evaluativa de carácter interdisciplinar, como un importante medio para facilitar fortalecer y extender la práctica evaluadora.

Pero la mejor demostración de que el enfoque transdisciplinar de la investigación evaluativa es algo totalmente asumido hoy en día, son los enfoques y contenidos que se publican en las revistas científicas promovidas por las Asociaciones Nacionales e Internacionales de Evaluadores y en todas las revistas relevantes en evaluación. A modo de ejemplo podemos citar las revistas “*African Journal of Evaluation*”, “*Canadian Journal of Program Evaluation*” y “*New Directions for Evaluation*”, promovidas por las Asociaciones de Evaluación africana, canadiense y americana respectivamente. También podemos citar a revistas de gran relevancia internacional, como son “*Evaluation and Program Planning*” y “*Evaluation*”, que señalan de manera expresa en su línea editorial el carácter interdisciplinar de la misma.

Como es lógico, las revistas científicas que se fundamentan en la transversalidad de la evaluación como disciplina, que son muchas más que las anteriormente aludidas a modo de ejemplos, conviven perfectamente con otras muchas que se centran en evaluación en ámbitos concretos de gran relevancia social como son la educación, la sanidad, la economía, las políticas sociales, etc., lo que también refuerza el carácter de transdisciplinar con el que se entiende actualmente a la investigación evaluativa.

El auge de la evaluación institucional

Entre los hitos de reciente fortalecimiento de la investigación evaluativa no podemos dejar de referirnos al auge tomado en las últimas décadas por la *evaluación institucional*, de manera muy acusada en el ámbito de la educación en sus instituciones en todos los niveles de carácter obligatorio y en la gran mayoría de los sistemas universitarios desarrollados en el contexto internacional (Escudero, 2007). Tomando como referencia inicial a las evaluaciones norteamericanas, este proceso se ha ido extendiendo paulatinamente en la mayoría de los sistemas universitarios, dentro de sus características específicas, de manera muy acusada en los ámbitos europeo y latinoamericano (Aguilar, 2001).

Un buen ejemplo del auge de la investigación evaluativa en este ámbito lo encontramos en la XIII Asamblea General de ALAFEC (Asociación Latinoamericana de Facultades y Escuelas de Contaduría y Administración) celebrada en Buenos Aires en Octubre de 2012, en la que se presentó como ponencia un análisis de la investigación evaluativa y su perspectiva en la evaluación de las instituciones de educación superior (Mira et al., 2012).

La evaluación institucional sintetiza y se centra en la gran mayoría de objetos de interés y de ámbitos de actuación de la investigación evaluativa, políticas, programas, sistemas organizativos y de gobierno, relaciones internas y externas, recursos personales y materiales, infraestructuras, etc., recogiendo e integrando además los principios y métodos defendidos desde el campo de la calidad institucional y de los procesos de su control, aseguramiento y mejora.

Esta realidad ha impulsado y fortalecido la utilización de ciertas metodologías en este campo de investigación tales como, por ejemplo, el desarrollo de auto-informes, la utilización de indicadores de rendimiento y los contrastes entre visiones internas y externas (Escudero, 2002). Desde la evaluación de instituciones y organizaciones, no sólo educativas sino también de otros ámbitos como, por ejemplo, el de la salud (Yusa, Hynie & Mitchell, 2016), se ha potenciado

extraordinariamente en la investigación evaluativa la utilización de evaluaciones internas y externas, como instrumentos obligados y complementarios para asegurar el acceso a todo tipo de información relevante, para llevar a cabo un diagnóstico completo de problemas por resolver y elementos por mejorar y reforzar la credibilidad y aceptación de la evaluación y sus propuestas, de cara a la toma de decisiones de mejora. Pérez Juste (2002) también resalta la importancia de ambos enfoques, el interno y el externo, para potenciar la calidad de la educación desde la evaluación de sus programas.

Un magnífico ejemplo de este panorama es el que ofrece una reciente publicación de la Comisión Europea (Eurydice 2015) en el que se analiza cómo evalúan la calidad de sus escuelas treinta y dos países europeos, comparando sus distintos enfoques y estructuras, e incidiendo de manera especial en los sistemas de evaluación interna y externa, sus procedimientos y tipo de evaluadores y la coordinación en la utilización de sus resultados.

Centrándose en el ámbito de la educación superior, Nicoletti (2013) nos ofrece un interesante análisis de la situación actual de este tipo de investigación evaluativa, de sus objetivos y de sus principales soportes metodológicos. Así mismo, en una reciente tesis doctoral en la Universidad de Zaragoza, el profesor Javier Paricio (2015) nos ofrece un completo análisis comparado de las concepciones de calidad y sistemas de garantía y aseguramiento de la misma que existen en el conjunto de las universidades europeas, diseñando al mismo tiempo un modelo metodológico de análisis multidimensional.

Al hilo de este impulso que dejan claro las referencias citadas, se han potenciado los enfoques y organismos de evaluación externa, creándose comisiones, agencias, oficinas intermedias y redes de carácter regional, nacional e internacional de apoyo a las evaluaciones y procesos de acreditación y certificación de calidad. Claverie, González y Pérez (2008), por ejemplo, nos describen y analizan críticamente el modelo de evaluación

de la calidad de la educación superior en Argentina, sustentado por la coordinación de la CONEAU (Comisión Nacional de Evaluación y Acreditación Universitaria). Reina Ferrández (2008) nos ofrece un análisis comparado de los programas de auditoría institucional universitaria en España y Gran Bretaña, resaltando la incidencia de organizaciones como la Agenda Nacional de Evaluación de la Calidad y Acreditación (ANECA) española y de la británica QAA (Quality Assurance Agency for Higher Education).

La diversidad y flexibilidad metodológica

Al igual que en las aproximaciones metodológicas, los enfoques de la investigación evaluativa pueden ser muy diversos, incluso con planteamientos evaluativos libres de objetivos, que, aunque no son muy habituales (Youker, Ingraham & Bayer, 2014), sí que pueden ser especialmente valiosos en determinados procesos de evaluación externa.

Expósito, Olmedo y Fernández-Cano (2004) nos ofrecen un ejemplo de esta diversidad en su revisión de las metodologías empleadas en la evaluación de programas en España. Christie y Fleischer (2010) también dejan constancia de este variado panorama en su análisis de publicaciones en revistas especializadas americanas.

La diversidad y flexibilidad en la metodología empleada en la investigación evaluativa es algo obligado por la propia naturaleza transdisciplinar y fundamentación paradigmática de este tipo de investigación, analizada en un apartado anterior, y por la multiplicidad de campos de trabajo, objetivos, objetos y contextos de análisis con los que el evaluador se ve obligado a trabajar, para elaborar sus diversos diagnósticos y propuestas de acción para los diferentes implicados, responsables y afectados por la investigación.

Esta pluri-realidad hace que la gran mayoría de las metodologías y procedimientos de investigación tengan utilidad para la

investigación evaluativa, pero, lógicamente, su aplicabilidad está condicionada y definida en cada caso por el tipo de problema, aspecto o temática que se está investigando. Lo mismo cabe decir de los procedimientos y fuentes de recogida de información, aspecto clave en muchos estudios evaluativos, que deben apoyarse en variables, informaciones, opiniones, etc., muy diversas en cuanto a su tipología, diversidad de implicados y posibilidades de recogida. Esta complejidad a la que se enfrenta habitualmente el evaluador, hace que la triangulación de métodos, técnicas, fuentes y procedimientos de recogida de información y la utilización combinada de distintas metodologías (*mixing methods*) adquiera una especial relevancia en la investigación evaluativa, como medio de fortalecimiento y validación de los diagnósticos y resultados (Betzner, Lawrenz & Thao, 2016; Makrakis & Kostoulas-Makrakis, 2016; Perassi, 2009; White, 2013).

En los últimos años se ha investigado bastante sobre la recogida de datos en sondeos a través de internet, algo que, naturalmente, ha abierto nuevas e interesantes perspectivas para los evaluadores. El número de otoño de 2007 de la revista *New Directions for Evaluation*, por ejemplo, está dedicado monográficamente a trabajos sobre la utilización de sondeos (online) en evaluación, incidiendo sobre sus fortalezas y limitaciones, así como en las condiciones que favorecen su utilización.

Esta apertura y adaptabilidad metodológica de la que estamos hablando se constata perfectamente en las publicaciones de estudios evaluativos en las revistas científicas especializadas, pues es muy habitual que ante la evaluación de temas, programas, planes, etc., novedosos o complejos, el equipo investigador diseñe, establezca y describa previamente un modelo y unos procedimientos metodológicos adaptados a su problema o tema específico, que aplica después en el desarrollo de su proyecto de evaluación.

Estrategias participativas

Desde hace varias décadas, y por múltiples y variadas razones tanto teóricas como

prácticas, los expertos han venido preocupándose y defendiendo la implicación de los afectados (*stakeholders*) en los propios programas de evaluación (Rodríguez-Campos, 2012). Una de esas razones es la de potenciar la utilización de los resultados de la evaluación en los distintos procesos de toma de decisiones que pueden afectar a los diversos implicados (Johnson et al. 2009; Patton, 2012). La idea básica es que si se participa y colabora con la evaluación, es mucho más fácil asumir sus resultados y sugerencias de acción y, consecuentemente, aceptar y promover los cambios necesarios, aunque sean costosos y supongan algunos sacrificios. Con la participación de los afectados, la evaluación, globalmente, aumenta su viabilidad y utilidad potencial. Además, se refuerza el principio de corresponsabilidad entre los implicados ante los problemas detectados y sus posibles decisiones y se potencian las fórmulas de gestión horizontal en los programas e instituciones.

Una de las claves para que este tipo de participación se produzca es que exista en las instituciones una cultura, esto es, un conjunto de supuestos básicos, normas, valores y visiones culturales compartidos por sus miembros, que influyan en su funcionamiento, integrando lo que podemos denominar como *cultura evaluativa* (Scheerens, 2004). En otras palabras, es necesaria la creencia de que la organización aprende y mejora a partir de la reflexión crítica y evaluación sistemática que llevan a cabo sus propios miembros. Una de las maneras más eficaces para vencer muchas de las inhibiciones de, por ejemplo, los profesores ante la evaluación, son los procesos de reflexión crítica sobre su propia tarea y el funcionamiento institucional.

Las fórmulas de colaboración y participación en las evaluaciones son múltiples y variadas en cuanto a intensidad e implicación (Taut, 2008), aunque todas ellas se sustentan de alguna manera en modelos organizativos reconocidos genéricamente como democráticos. En algunas ocasiones con procesos de una sola fase, en la que todos participan, y en otras con varias fases, con

distintos grados de implicación de unos y otros. Askew y colaboradores (2012) defienden la necesidad de compatibilizar las técnicas colaborativas en las evaluaciones con enfoques respondientes a las culturas diversas de los implicados, para evitar la discriminación de grupos minoritarios. Geist (2010), por ejemplo, plantea y explica la utilización del método Delphi como procedimiento para impulsar la participación de los implicados (stakeholders) en el proceso evaluativo. Rodríguez-Campos (2012a) ofrece una interesante revisión de los avances recientes de la evaluación colaborativa y de su fundamentación, con sus fortalezas y aportaciones, planteando, además, un enfoque modélico de utilización de la misma. Al mismo tiempo, se sigue investigando empíricamente sobre la participación en las evaluaciones y sobre la relación entre la implicación y uso en los procesos de evaluación y las fórmulas de colaboración más eficaces en los diferentes contextos y tipologías de la evaluación (Brandon & Fukunaga, 2014; Chouinard & Milley, 2016; Daigneault, 2014; Roseland et al., 2015).

La idea de evaluación democrática se ha ido desarrollando con diversos enfoques y aproximaciones (Ryan, 2004), que se reflejan en los aspectos organizativos y metodológicos, en la generación del conocimiento, en el papel de los evaluadores y en las prácticas de comunicación entre los implicados. Una fórmula extraordinariamente extendida en la evaluación institucional, por ejemplo, es la de conjugar fases de auto-evaluación y evaluación externa, buscando la integración de los resultados y la complementariedad de las posiciones e intereses de las distintas audiencias. Algunos autores van todavía más lejos (Christie y otros, 2004), pues defienden que ambas evaluaciones sean llevadas a cabo por un único equipo, compuesto de agentes internos y externos, trabajando en todo momento en colaboración.

En el ámbito de la salud, por ejemplo, de cara a la toma de decisiones, Adelson y colaboradores (2003), tras una amplia revisión del tema, resaltan la importancia y utilidad del

uso de metodologías deliberativas, con la participación de diversos grupos de implicados. También en este ámbito, un grupo de investigadores austriacos, pertenecientes al Instituto Ludwig Boltzmann de Investigación en Promoción de la Salud en Viena (Nitsch et al, 2013) ofrecen una interesante revisión de casi medio centenar de artículos científicos de evaluación en programas de promoción de salud, analizando los modelos y niveles de participación de los implicados (stakeholders), así como las fortalezas y debilidades de tales procesos.

El fortalecimiento de las estrategias colaborativas en los últimos años, en la práctica de la investigación evaluativa, es incuestionable, convirtiéndose en el momento presente en algo de uso habitual en los proyectos de evaluación (Preskill & Boyle, 2008). El gran mentor de la evaluación para el empoderamiento (Fetterman, 2001a), nos viene a decir que en el siglo XXI, la evaluación se ha convertido en un proceso de colaboración entre los implicados en la misma.

Fitzpatrick (2012) ofrece una revisión de cómo se ha ido extendiendo el uso de la evaluación colaborativa desde el siglo pasado, así como de los diferentes modelos y aproximaciones de colaboración recogidos en la literatura científica, de sus fortalezas y limitaciones, y de cómo su aplicación está condicionada por la naturaleza del problema evaluado, el contexto en el que tiene lugar y la tipología de los implicados.

Utilidad y uso de la evaluación

El análisis de la literatura científica nos confirma que, en el momento presente, en pocas cosas están tan de acuerdo los expertos en investigación evaluativa como en el hecho de que esta se lleva a cabo con la intención de ser útil e influir en la toma de decisiones posterior. De hecho, los propios evaluadores dedican grandes esfuerzos a conseguir que sus sugerencias sean tenidas en cuenta y aplicadas convenientemente (Christie, 2007). Sin embargo, la utilización de los resultados de las evaluaciones, por múltiples razones, siempre

es un asunto sometido a controversias (Escudero, 2013).

Para fortalecer una adecuada utilización de las evaluaciones, es importante analizar bien los contextos de decisión de los usuarios potenciales y centrar el estudio en aspectos relevantes para dicha tarea (Ledermann, 2012; Leviton, 2003; Neuman et al., 2013). El desarrollo de la cultura evaluativa en las organizaciones es otro elemento facilitador de la utilización adecuada de los procesos evaluativos (Cousins et al. 2014).

Reflexionando desde la práctica, Grasso (2003) nos dice que, de antemano, nunca vamos a tener completa seguridad de que una evaluación se utilizará, pero que esta posibilidad se acrecienta si se optimizan las condiciones para ello. En concreto, una evaluación es más *útil* y más *utilizable*, si el evaluador se ha respondido previamente a las tres preguntas siguientes:

- ¿Quiénes utilizarán la evaluación?
- ¿Qué necesitarán de la evaluación?
- ¿Cuándo necesitarán la información?

En definitiva, parece necesario determinar cuáles son las *audiencias*, cómo se les puede hacer llegar la información, cuáles son las prioridades de cada una de ellas y con qué celeridad se les debe atender. Por todo ello, es fundamental el contacto y las consultas con las diversas audiencias antes de diseñar y llevar a cabo la evaluación, para favorecer una comunicación comprensible y en el momento apropiado de los resultados de la evaluación y de las sugerencias de acción. Hoy en día, partiendo de este tipo de principios, se desarrollan y aplican distintos sistemas de auditoría de la información producida por las evaluaciones (Schwarth & Mayne, 2005), para asegurar su calidad y, al mismo tiempo, que esté disponible cuando sea preciso, que responda a las demandas del problema y que sea comprensible para los que la necesitan y demandan.

También existe una corriente de evaluadores (Henry & Mark, 2003; Kirkhart, 2000) que defienden que el concepto de uso de

la evaluación, entendido como utilización directa y focalizada, se ha quedado estrecho, por lo que debe darse paso al concepto de *influencia*, como idea de consecuencias más abiertas y amplias, y no necesariamente delimitadas de antemano. La influencia habla también de las consecuencias posibles y no sólo de las buscadas, esto es, de la capacidad de producir cambios por medios intangibles o indirectos.

Profundizando en la anterior línea de pensamiento, Henry (2003) enfatiza el gran valor de la evaluación en las democracias, como instrumento de mejora social, aunque, según su opinión, los evaluadores no han sabido explotar todo su potencial para influir en este sentido, probablemente porque la investigación sobre el uso de la evaluación se ha centrado en su utilización inmediata.

En la práctica, algo que condiciona extraordinariamente el uso o influencia de la evaluación es la manera en que se comunica o se informa de ella. Es importante resaltar que la buena investigación evaluativa no sólo se fundamenta en buenas técnicas analíticas, como cualquier otro tipo de investigación social, pues debe combinar éstas con técnicas de comunicación bien desarrolladas. En otras palabras, no basta con que la investigación evaluativa esté bien hecha, además, debe estar bien contada, no sólo a los especialistas, sino a todos los implicados y afectados por la misma, a todas las audiencias. A esto hay que añadir que hay especialistas (Perrin, 2001), que ponen mucho énfasis en la conveniencia de que los evaluadores se apoyen siempre en los métodos más simples de entre los que permitan responder a las preguntas planteadas, pues esto no sólo tiene consecuencias en los costes, rapidez y aceptación de la evaluación, sino también en su uso.

Por lo anterior, existen distintos planteamientos sobre cómo se debe informar de las evaluaciones con vistas a *optimizar su visibilidad* y, consecuentemente, favorecer su utilización. A partir de un informe común del proceso de evaluación y sus resultados, se pueden introducir sub-apartados, apéndices, sugerencias específicas, etc., dirigidos a las

distintas audiencias implicadas. Otra opción es la de elaborar informes separados según las audiencias, aunque esta es una fórmula menos habitual.

Lo que sí resulta crítico en todo caso, es la adecuación del lenguaje técnico a la hora de presentar los resultados que, en la investigación evaluativa, las más de las veces son producto de algún tipo de tratamiento estadístico, que no es familiar para muchos de los potenciales receptores de tales resultados. El evaluador debe ser capaz de presentar sus resultados apoyándose en la jerga estadística convencional y, además, de hacerlo de forma que tenga significado (May, 2004) para las distintas audiencias.

Este autor nos dice que un resultado estadístico tiene significado para una audiencia determinada, cuando es:

- *Comprensible*, esto es, que se entienda de manera fácil por la mayoría de la gente, con el mínimo de supuestos posibles sobre la preparación estadística de la audiencia y evitando la jerga estadística.
- *Interpretable*, esto es, utilizando unidades de medida familiares o fácilmente explicables.
- *Comparable*, esto es, que permita observar las diferencias con otras situaciones, a lo largo del tiempo, otros programas, otros centros, etc.

El tercer criterio, el de la comparabilidad, nos dice May que es menos crítico que los dos primeros, pues no siempre se dan las condiciones de necesidad o propicias para la comparación. Cuando esto sucede, el criterio enriquece mucho el valor de los anteriores.

Ética y utilización de la evaluación

Hace unos años, House (2008) lanzó una llamada de atención contra el peligro de infectar y poner en serio peligro el campo de la investigación evaluativa con malas prácticas procedentes de un terreno con una buena imagen en cuanto al diseño de sus evaluaciones. Se refería en concreto a los procesos de evaluación de fármacos que, según este autor, con cierta frecuencia están

sesgados a enfatizar excesivamente los efectos positivos del fármaco evaluado y a esconder los negativos, e incluso manipulando los diseños de investigación para obtener los resultados deseados. No se puede dejar sólo en manos de las empresas beneficiadas, ni del mercado, señala este autor, un asunto tan crítico e importante desde el punto de vista social. En este y en todos los asuntos sociales, los responsables públicos y los propios evaluadores deben luchar frontalmente contra las pseudo-evaluaciones y contra las malas evaluaciones y las evaluaciones manipuladas (misevaluations). También el contexto político en el que se desarrolla una evaluación puede ser condicionante de la misma y de la consiguiente toma de decisiones, por ello, el fortalecer la independencia de los evaluadores sigue siendo un tema de preocupación y estudio para los especialistas en investigación evaluativa (Azzam & Levine, 2015).

Como nos señalan Pinkerton et al. (2002), el tema ético es crítico en los estudios de coste-eficacia, coste-beneficio y coste-utilidad de los programas sociales de gran alcance, pues entran en escena decisiones específicas de base ideológica, que condicionan la aportación económica con una orientación u otra, beneficiando a unos implicados u otros.

El mismo evaluador puede ser una fuente de conflictos cuando sus valores personales son muy contrarios, o son muy favorables, hacia el objeto mismo de la evaluación. Así, por ejemplo, alguien implicado personalmente en la defensa o en la crítica de un programa educativo concreto, es muy difícil que pueda evaluarlo sin generar dudas sobre su trabajo.

Finalmente, el propio objeto de la evaluación puede ser conflictivo y le puede plantear problemas de naturaleza ética al investigador. Un caso habitual es cuando existen defensores y detractores del programa evaluado. Entonces, el evaluador debe dar audiencia a todas las partes en conflicto y articular un sistema armonioso de confluencia de criterios y de elaboración de conclusiones y sugerencias (Abma, 2000).

Un segundo tipo de situación conflictiva se produce cuando se evalúan intervenciones sociales que no tienen el soporte de la investigación. En tal caso, el evaluador debe ser extremadamente riguroso en el análisis de resultados de la misma, contrastando los efectos en diferentes contextos y comparando tal realidad con la que se da en situaciones sin intervención. En suma, en esta circunstancia el evaluador debe dejar constancia de su condición de investigador experimental.

Cruzando las dimensiones de (*uso-no uso*) y (*uso correcto-mal uso*), Cousins (2004) establece cuatro cuadrantes o categorías con relación a la utilización de las evaluaciones. Estas categorías son las siguientes:

- 1) *Utilización ideal*, en la que se recogen los usos instrumentales, conceptuales y de persuasión o apoyo antes comentados.
- 2) *Mala utilización* (“misuse”), bien por incompetencia o incomprensión de los resultados de la evaluación, bien por manipulación intencionada de los mismos.
- 3) *No utilización injustificada*, desconsiderando de manera abusiva o eliminando los resultados de la evaluación.
- 4) *No utilización justificada*, cuando aparecen razones de diverso tipo que aconsejan no tener en cuenta los resultados de la evaluación.

Aunque no resulte fácil, por su propia situación en el proceso de toma de decisiones, el evaluador no sólo debe intentar evitar que sus resultados no sean utilizados de manera injustificada (categoría 3), por desgracia una situación muy frecuente, sino que debe hacer todo lo posible por evitar el mal uso de su trabajo (categoría 2), sin duda el problema más peligroso.

Como es lógico, la investigación evaluativa que es verdaderamente útil y rigurosa, se hace al margen de estas situaciones que ponen en peligro la ética en las evaluaciones y en sus usos, algo que, como hemos analizado, es fundamental para el desarrollo y credibilidad de un tipo de

investigación aplicada cada vez más relevante para asegurar y acrecentar la calidad de nuestras políticas sociales en general y, en definitiva, de nuestra sociedad. Una investigación que, como señala Chelinsky (2008), debe buscar en todo momento el ajuste entre la necesaria independencia científica y las justas demandas de una sociedad democrática.

Por estos motivos, las asociaciones profesionales de evaluadores han elaborado documentos donde establecen los principios éticos para el desarrollo de las evaluaciones. Un ejemplo es la “*Guiding Principles for Evaluators*” de la AEA (American Evaluation Association, 2008), orientada para promover la práctica ética en la evaluación de programas, productos, personal y políticas. Los principios tratados son la indagación sistemática, la competencia, la integridad/honestidad, el respeto a las personas y la responsabilidad con el bienestar general y de los ciudadanos.

Schweigert (2007) indaga también en este terreno de la ética, indicando que la justicia es una prioridad en la acción evaluadora, que debe preservarse con la colaboración de los evaluadores y el conjunto de implicados y afectados. Este autor nos habla de tres dimensiones en la justicia a preservar, en primer lugar la que denomina *justicia pública*, como las obligaciones, principios y normas de la profesión de evaluador, en segundo lugar la *justicia procedimental*, asegurando el respeto a los derechos y el trato adecuado a todos los implicados y, finalmente, la *justicia distributiva*, a la búsqueda de un desarrollo social equitativo.

Metaevaluación y normas de calidad

La metaevaluación, *la evaluación de la propia evaluación*, adquiere cada vez más importancia como consecuencia del desarrollo de la investigación evaluativa y de la expansión de proyectos de evaluación en todos los terrenos. Parece lógico pensar que lo primero que se necesita es *confianza* en algo que, precisamente, se hace para *tener confianza* a la hora de tomar decisiones, intervenir o sugerir algo sobre un proyecto

social. Se trata de articular la obligación científica de reflexionar sobre lo que puede estar yendo bien o mal con un plan determinado (Escudero, Pino & Rodríguez, 2010).

Un autor de larga tradición en el tratamiento con el tema como Stufflebeam (2001b), además de indicar que la metaevaluación debe juzgar el mérito, el valor y la utilidad de la evaluación, la establece como elemento de garantía científico-profesional; los propios evaluadores deben asegurar que las investigaciones evaluativas estén bien hechas, que se mejoren sistemáticamente y que den garantía a sus audiencias. Por ello, la metaevaluación es una obligación profesional de los evaluadores, que entronca con la ética de la revisión y el autocuestionamiento continuos (Stake, 2006). La metaevaluación intenta detectar sesgos de tipo estadístico, desviaciones en la interpretación de los datos, problemas de comprensión de algunos aspectos, etc., dentro de una perspectiva crítica que parte de la idea de que toda evaluación es mejorable

Siguiendo al propio Stufflebeam (2001), podemos definir metaevaluación como «el proceso de diseño, obtención y aplicación de información descriptiva y valorativa- acerca de la utilidad, factibilidad, legitimidad y rigor de una evaluación, así como de su sistematicidad y de la competencia, integridad, honestidad, respetabilidad y responsabilidad social de sus autores- sobre el desarrollo de la evaluación e informar públicamente de sus fortalezas y debilidades» (185). Este mismo autor nos dice que un buen símil para entender este concepto es el de que la *metaevaluación* es en el campo de la *evaluación* lo que la *auditoría* es en el campo de la *contabilidad*, que asegura el escrutinio externo e independiente del uso razonable de los recursos financieros.

En la metaevaluación también confluyen las perspectivas *formativa* y *sumativa*. En la primera, como una guía de ayuda a los evaluadores a preparar todo, antes de realizar la evaluación, en la metaevaluación sumativa se determina el mérito, valor y utilidad de la evaluación una vez terminada. Ambas

perspectivas conjuntamente determinan el grado en que la evaluación es técnicamente adecuada, útil para la toma de decisiones, ética en su incidencia sobre personas y eficiente en el uso de recursos. En la aproximación formativa el metaevaluador actúa como un *técnico consultor* y en la sumativa como un *agente externo certificador*.

El trabajo metaevaluador puede llevarse a cabo según distintos enfoques modélicos (Escudero, 2011). Así por ejemplo, *la revisión del estudio* se basa fundamentalmente en el análisis del informe de evaluación, de los procesos y resultados; es un procedimiento esencial en el estudio de naturaleza interpretativa, fundamentalmente sobre datos cualitativos. *El análisis secundario* supone el reanálisis de los datos originales, siendo un procedimiento muy poderoso si se puede llevar a cabo correctamente, pues el hecho de disponer de los datos de una evaluación ya es signo de credibilidad de la misma. *La integración cuantitativa* o *meta-análisis* es un procedimiento estadístico de combinación de resultados cuando existen varios estudios sobre el mismo tema. También podemos hablar de metaevaluaciones internas y externas.

Las metaevaluaciones se apoyan en técnicas muy variadas, aunque existen algunas de creciente utilización como los paneles de revisión, análisis de protocolos, seminarios interactivos y las listas de verificación (checklists) adecuadas a evaluaciones de distinta naturaleza y contenido (Scriven, 2000; Stufflebeam, 2000). Ahondando en estas ideas, Vanhoof y Van Petegen (2010) nos indican que la metaevaluación requiere múltiples criterios, múltiples métodos y múltiples personas.

Para juzgar la calidad de las evaluaciones se han ofrecido diversos criterios, en general ligados con el mérito y el valor, la utilidad, la eficiencia del diseño, etc. En general, se habla de criterios de: a) *rigor* (métodos precisos correctamente aplicados), b) *utilidad* (valor, mérito, importancia de resultados) y c) *eficiencia y efectividad* del diseño y del proceso evaluador. También se habla de

criterios *éticos*, en general, como protección de los derechos de los implicados en la evaluación, preservando la verdad de lo evaluado, en respuesta a determinadas presiones distorsionadoras.

En relación con el desarrollo de la metaevaluación, en un contexto de crecimiento profesional de la investigación evaluativa, está el establecimiento de *normas* bajo las que desarrollar la evaluación con vistas a asegurar su calidad y evitar el máximo de potenciales defectos. Así, por ejemplo, durante las últimas décadas se han elaborado normas sobre investigación y medición educativas por distintas asociaciones científicas tales como la *American Psychological Association* (APA), la *American Educational Research Association* (AERA) y el *National Council on Measurement in Education* (NCME).

Pero a nadie se le escapa que el prestigio académico y social que la investigación evaluativa como disciplina y la tarea de los evaluadores profesionales han venido adquiriendo y aumentando en las últimas décadas, está en buena medida marcado por el liderazgo y tutela del “*Joint Committee on Standards for Educational Evaluation*”, creado por las tres asociaciones científicas antes citadas, que empezó a trabajar en 1975. Este Comité está actualmente auspiciado por más de una quincena de organizaciones académico-profesionales, fundamentalmente americanas y canadienses, del ámbito de la evaluación educativa (Joint Committee, 2003), y su mayor aportación ha sido la elaboración y el establecimiento de unas *normas ético-científicas*, de unos estándares, para regular la evaluación educativa en Estados Unidos y Canadá. Sin embargo, su aplicación ha ido mucho más allá del aludido espacio geográfico y del ámbito de la educación, aunque se hayan suscitado algunos problemas de adaptación en determinados contextos (Stufflebeam, 2004).

En el ámbito africano, por ejemplo, la “*African Evaluation Associations and Networks*” ha adaptado estas normas, elaborando un conjunto de 30 criterios de control de calidad que denominan “*African Evaluation Guidelines*” (Patel, 2002a, 2002b).

El objetivo básico del Comité Conjunto y de sus normas, que se entienden como algo dinámico y que debe revisarse y adecuarse periódicamente, es el de promover conciencia sobre la necesidad de que las evaluaciones sean sólidas, de calidad, útiles y que respondan a las necesidades de evaluación. Para conseguir esto, el Comité establece normas de los cuatro tipos siguientes: de *utilidad*, de *factibilidad*, de *legitimidad* y de *precisión*.

Es muy importante observar el orden en el que se plantean las distintas normas, pues actúan como llave o condición previa para las siguientes. En concreto, al plantearse un proyecto evaluativo, la primera condición que hay que asegurar es que sea *útil*, esto es, que tenga sentido llevarlo a cabo para mejorar el sistema, tomar decisiones sumativas, planificar, etc. En otras palabras, el Comité Conjunto nos viene a sugerir en alguna medida que, de entrada, una evaluación tiene que ser *útil*, y que si esto no es así, no tiene sentido plantearla.

Asegurada la utilidad de una evaluación, es cuando hay que analizar si es *viable*, si se puede hacer, después si es *legítima*, si no viola principios de legitimidad, derechos de implicados, etc., y, finalmente, se analiza, diseña y se realiza con el máximo *rigor científico*, siguiendo los principios y la metodología más adecuados y poderosos.

Hoy en día, las bases para este proceso metaevaluativo, no olvidando que el contexto natural de su definición son los Estados Unidos y Canadá, con lo que su aplicación en otros contextos requiere contextualización (Jang, 2000), están en los aludidos estándares de evaluación del *Joint Committee* y en los principios guía de la *American Evaluation Association*, definidos para todos los ámbitos y no sólo el educativo. Sin estos principios ético-científicos reguladores del trabajo de los evaluadores, es indudable que la investigación evaluativa no habría alcanzado el prestigio y la relevancia que tiene en el momento presente (Escudero 2011).

Parece indudable que la gran importancia e creciente influencia que la investigación

evaluativa ha tenido y tiene en las últimas décadas en la educación y en otros ámbitos sociales, habría sido más cuestionada si las asociaciones profesionales más directamente implicadas no hubieran establecido estas *normas ético-científicas*, para regular el trabajo de los evaluadores.

Reflexión para finalizar

En este trabajo presentamos y valoramos la consolidación de la investigación evaluativa como una disciplina transversal de gran pujanza, especialmente en el campo de las ciencias sociales, así como su creciente utilización, como campo de investigación, en múltiples ámbitos del conocimiento y del desarrollo social.

Partiendo de la definición actual de investigación evaluativa, hemos enfatizado la función de esta al servicio del cambio social y, en concreto, al servicio de la mejora social. Hemos visto que la evaluación resulta ser un elemento imprescindible de planificación, de gestión y de toma de decisiones, con el gran objetivo de mejorar la sociedad y, más en concreto, la calidad de vida de los ciudadanos, pero sabemos que hablamos de algo complejo, dinámico, multidimensional, necesitado de contextualización, que requiere análisis y tratamientos rigurosos y reposados.

El análisis de la utilidad y el uso de la evaluación constituye un bloque central en este trabajo porque se trata de uno de los mayores problemas con el que hoy se encuentran los evaluadores como profesionales, porque no siempre se les presta la debida atención a los resultados y sugerencias de sus trabajos, y porque el futuro de la investigación evaluativa como disciplina académica y como profesión relevante, está directamente relacionado, y condicionado, al buen uso y a la correcta utilización que se haga de las evaluaciones y de sus resultados. Si las investigaciones evaluativas no son útiles para el cambio y la mejora social y si no se utilizan bien, nuestra tesis es que son un esfuerzo poco relevante y, por lo tanto, poco justificable. Pero sabemos y defendemos que si se utilizan bien, son una herramienta de especial relevancia estratégica.

En el trabajo presentamos distintos modelos de utilización, usos de los resultados, así como distintas fórmulas y estrategias facilitadoras de esta utilización. Entramos también en el debate que ha surgido recientemente alrededor de los conceptos de uso y de influencia. También hacemos una defensa nítida de las estrategias participativas en la evaluación como estímulo y garantía de uso y buena utilización de sus resultados, señalando las distintas fórmulas de participación que se plantean o pueden plantearse.

El trabajo aporta también una llamada de atención sobre la utilización ética de las evaluaciones, con la presentación de los principios que deben seguirse, y se resalta, asimismo, la relevancia de las metaevaluaciones como instrumentos de garantía de calidad de las investigaciones evaluativas.

Como idea final, queremos señalar la creciente importancia que la investigación evaluativa tiene en todos los ámbitos científicos, departamentos universitarios, institutos y otros organismos de investigación, etc. conviviendo con otros enfoques investigadores de investigación básica y aplicada, tanto en el terreno de la formación como en el desarrollo de proyectos.

Referencias

- Abelson, J., Forest, P-G., Eyles, J., Smith, P., Martin, E., & Gauvin, F.P. (2003). Deliberations about deliberative methods: issues in the design and evaluation of public participation processes. *Social Science & Medicine* 57, 239–251.
- Abma, T. A. (2000). Stakeholder conflict: a case study. *Evaluation and Program Planning* 23, 199-210.
- Aguilar, M. (2001). La evaluación institucional de las universidades. Tendencias y desafíos. *Revista de Ciencias Sociales* (Cr), II-III, 93-92, 23-34.
- American Evaluation Association (2008). Guiding Principles for Evaluators *American Journal of Evaluation*, 29 (4), 397-398.

- Askew, K., Green Beverly, M. & Jay. M. L. (2012). Aligning collaborative and culturally responsive evaluation approaches. *Evaluation and Program Planning*, 35, 552–557.
- Azzam, T. & Levine, B. (2015). Politics in evaluation: Politically responsive evaluation in high stakes environments *Evaluation and Program Planning*, 53, 44–56. DOI: <http://dx.doi.org/10.1016/j.evalprogplan.2015.07.002>
- Betzner, A., Lawrenz, F. P. & Thao. M. (2016). Examining mixing methods in an evaluation of a smoking cessation program. *Evaluation and Program Planning*, 54, 94–101. DOI: <http://dx.doi.org/10.1016/j.evalprogplan.2015.06.004>
- Brandon, P. R. & Fukunaga, L. L. (2014). The state of the empirical research literature on stakeholder involvement in program evaluation. *American Journal of Evaluation*, 35 (1), 26–44. DOI: <http://dx.doi.org/10.1177/1098214013503699>
- Bredo, E. (2006). Philosophies of Educational Research. En J. L. Green, G. Camilli & P. B. Elmore, *Handbook of Complementary Methods in Education Research*. London: Lawrence Erlbaum Associates Publishers AERA 3-31.
- Calderon, A. J. (2004). Institutional Research at RMIT. A case study. Ponencia presentada en el 26th EAIR Forum, Barcelona, 5-8 de septiembre.
- Chelimsy, E. (2008). A Clash of Cultures: Improving the “Fit” Between Evaluative Independence and the Political Requirements of a Democratic Society. *American Journal of Evaluation*, 29, 4, 400-415.
- Chouinard, J. A., & Milley. P. (2016). Mapping the spatial dimensions of participatory practice: A discussion of context in evaluation. *Evaluation and Program Planning*, 54, 1-10. <http://dx.doi.org/10.1016/j.evalprogplan.2015.09.003>
- Christie, C. A. (2003). The practice-theory relationship in evaluation. *New Directions for Program Evaluation*, 97, Jossey-Bass, San Francisco, Ca.
- Christie, C. A. (2007). Reported influence of evaluation data on decision makers’ actions: An empirical examination. *American Journal of Evaluation*, 28, 1, 8–25
- Christie, C. A. & Fleischer, D. N. (2010). Insight Into Evaluation Practice: A Content Analysis of Designs and Methods Used in Evaluation Studies Published in North American Evaluation-Focused Journals. *American Journal of Evaluation*, 31(3), 326-346.
- Christie, C. A., Ross, R. M. & Klein, B. M. (2004). Moving toward collaboration by creating a participatory internal-external evaluation team: a case study. *Studies in Educational Evaluation*, 36(2), 107-117.
- Claverie, J., Gonzalez, G. & Perez, L. (2008), El Sistema de Evaluación de la Calidad de la Educación Superior en la Argentina: El Modelo de la CONEAU. Alcances y Límites para Pensar la Mejora. *Revista Iberoamericana de Evaluación Educativa*, 1(2), 149-164.
- Cook, J. R. (2015). Using Evaluation to Effect Social Change: Looking Through a Community Psychology Lens. *American Journal of Evaluation*, 28(1), 107-117. DOI: <http://dx.doi.org/10.1177/1098214014558504>
- Cousins, J. B. (2004). Commentary: Minimizing evaluation misuse as principled practice. *American Journal of Evaluation*, 25(3), 391-397.
- Cousins, J. B.; Goh, S.C.; Elliot, C.J. & Bourgeois, I. (2014). Framing the Capacity to Do and Use Evaluation. *New Directions for Evaluation*, 141 (1), 7-23. DOI: <http://dx.doi.org/10.1002/ev.20076>
- Daigneault, P. (2014). Taking stock of four decades of quantitative research on stakeholder participation and evaluation

- use: A systematic map. *Evaluation and Program Planning*, 45, 171–181. DOI: <http://dx.doi.org/10.1016/j.evalprogplan.2014.04.003>
- Donaldson, S. I. (2007). *Program theory-driven evaluation science*. Mahwah, NJ: Erlbaum
- Donaldson, S. I. & Gooler, L. E. (2003). Theory-driven evaluation in action: lessons from a \$20 million statewide Work and Health Initiative. *Evaluation and Program Planning*, 26, 355–366
- Donaldson, S. I. & Lipsey, M. W. (2006). Roles for theory in contemporary evaluation practice: Developing practical knowledge. En I. Shaw, J. C. Greene & M. M. Mark (Eds.), *The Handbook of Evaluation: Policies, programs, and Practices* (56-75). London: Sage
- Donaldson, S. I. & Scriven, M. (2003). Diverse visions for evaluation in the new millennium. En S. I. Donaldson & M. Scriven (Eds.) *Evaluating social programs and problems: Vision for the new millenium* (pp. 3-16), Mahwah, NJ: Erlbaum.
- Escudero, T. (2000). Evaluación de centros e instituciones educativas: las perspectivas del evaluador. En D. González, E. Hidalgo & J. Gutiérrez (Coords.), *Innovación en la escuela y mejora de la calidad educativa* (pp. 57-76). Granada: Grupo Editorial Universitario
- Escudero, T. (2002). Evaluación institucional: algunos fundamentos y razones. En V. Álvarez & A. Lázaro, *Calidad de las Universidades y Orientación Universitaria* (pp. 103-138). Málaga: Ediciones Aljibe.
- Escudero, T. (2003). Desde los tests hasta la investigación evaluativa actual. Un siglo, el XX, de intenso desarrollo de la evaluación en educación. *RELIEVE*, 9(1), DOI: <http://dx.doi.org/10.7203/relieve.9.1.4348>
- Escudero, T. (2005-2006). Claves identificativas de la investigación evaluativa: análisis desde la práctica. *Contextos Educativos. Revista de Educación*, 8-9, 179-199.
- Escudero, T. (2006). Evaluación y mejora de la calidad en educación. En T. Escudero & A. D. Correa, *Investigación en innovación educativa: algunos ámbitos relevantes* (pp. 269-325). Madrid: La Muralla, S. A.
- Escudero, T. (2007). Evaluación institucional de la calidad universitaria en España: Breve pero interesante historia. *Anuario de Pedagogía*, 9, 103-115.
- Escudero, T. (2009). Some relevant topics in educational evaluation research. En M. Asorey, J. V García Esteve, M. Rañada, & J. Sesma, *Mathematical Physics and Field Theory. Julio Abad, in Memoriam*, (pp. 223-230). Prensas Universitarias de Zaragoza.
- Escudero, T., Pino, J. L. & Rodríguez, C. (2010). Evaluación del profesorado universitario para incentivos individuales: revisión metaevaluativa. *Revista de Educación*, 351, 513-537.
- Escudero, T. (2011). *La construcción de la investigación evaluativa. El aporte desde la educación*. Prensas Universitarias-Universidad de Zaragoza.
- Escudero, T. (2013). Utilidad y uso de las evaluaciones. Un asunto relevante. *Revista de evaluación educativa*, 2 (1). Recuperado de <http://revalue.mx/revista/index.php/revalue/issue/current>
- European Commission/EACEA/EURYDICE (2015). *Assuring Quality in Education: Policy and Approaches to School Evaluation in Europe*, Eurydice Report. Luxembourg: Publication Office of the European Union. Recuperado de <http://eacea.ec.europa.eu/education/eurydice/>
- Exposito, J., Olmedo, E. & Fernandez-Cano, A. (2004). Patrones metodológicos en la investigación española sobre evaluación de programas educativos. *RELIEVE*, 10 (2). DOI: <http://dx.doi.org/10.7203/relieve.10.2.4319>
- Ferrandez, R. (2008). Programas de Auditoría Institucional Universitaria. Comparación de la Propuesta Española con el Sistema

- Británico. *Revista Iberoamericana de Evaluación Educativa*, (1)1, 156-170.
- Fetterman, D. M. (2001a). The Transformation of Evaluation into a Collaboration: A Vision of Evaluation in the 21st Century. *American Journal of Evaluation*, (22) 3, 381-385.
- Fetterman, D. M. (2001b). *Foundations of empowerment evaluation*. Thousand Oaks, CA.: Sage.
- Fetterman, D.M., Kaftarian, S.J. & Wandersman A. (Eds.) (2015). *Empowerment Evaluation: Knowledge and Tools for Self- Assessment, Evaluation Capacity Building, and Accountability*(2nd edition). Thousand Oaks, CA.: Sage Publications.
- Fitzpatrick, J. L. (2012). Commentary- Collaborative evaluation within the larger evaluation context, *Evaluation and Program Planning*, 35, 558-563-
- Geist, M. R. (2010). Using the Delphi method to engage stakeholders: A comparison of two studies. *Evaluation and Program Planning*, 33, 147-154.
- Henry, G. T. (2003). Influential evaluations. *American Journal of evaluation*, 24 (4), 515-524.
- Henry, G. T. & Mark, M. M. (2003). Beyond use: Understanding evaluation's influence on attitudes and actions, *American Journal of Evaluation*, 24 (3), 293-314.
- House, E. R. (2008). Blowback. Consequences of Evaluation for Evaluation. *American Journal of Evaluation*, 29(4), 416-426.
- Jacob, S. (2008). Cross-Disciplinarization: A New Talisman for Evaluation? *American Journal of Evaluation*, 29 (2), 175-194.
- Johson, K. (2009). Research on Evaluation Use: A Review of the Empirical Literature From 1986 to 2005. *American Journal of Evaluation*, 30(3), 377-410.
- Joint Committee On Standards Of Educational Evaluation (2003). *The student evaluation standards*. Thousand Oaks, Ca.: Corwin
- Kirkhart, K. (2000). Reconceptualizing evaluation use: An integrated theory of influence. En V. Caracelli, & H. Preskill (Eds.), *The expanding scope of evaluation use. New Directions for Evaluation*, N° 88, San Francisco, Ca.: Jossey-Bass.
- La Velle, J. M. & Donaldson, S. I. (2010). University-Based Evaluation Training Programs in the United States 1980-2008: An Empirical Examination. *American Journal of Evaluation*, 31(1), 9-23.
- Leviton, L. C. (2003). Evaluation use: Advances, challenges and applications. *American Journal of Evaluation*, 33 (2), 159-178.
- Ledermann, S. (2012). Exploring the Necessary Conditions for Evaluation Use in Program Change. *American Journal of Evaluation*, 24 (4), 525-535. DOI: <http://dx.doi.org/10.1177/1098214011411573>
- Makrakis, V. & Kostoulas-Makrakis, N. (2016). Bridging the qualitative–quantitative divide: Experiences from conducting a mixed methods evaluation in the RUCAS programme. *Evaluation and Program Planning* 54, 144-151. DOI: <http://dx.doi.org/10.1016/j.evalprogplan.2015.07.008>
- Mark, M. M. (2003). Toward a integrative view of the theory and practice of program and policy evaluation. En S. I. Donaldson & M. Scriven (Eds.) *Evaluating social programs and problems: Vision for the new millenium*, (pp. 183-204). Mahwah, NJ.: Erlbaum.
- Maxcy, S. J. (2003). Pragmatic threads in mixed methods research in social sciences: An emerging theory in support of practice. En A. Tashakkori & C. Teddlie (Eds.), *Handbook of mixed methods in social and behavioral research*, (pp. 51-89). Thousand Oaks, CA: Sage
- May, H. (2004). Making statistics more meaningful for policy research and program evaluation. *American Journal of Evaluation*, 25(4), 525-540.

- Mcclintock, C. (2003). Commentary: The evaluator as scholar/practitioner/change agent. *American Journal of Evaluation*, 24(1), 91-96.
- Mira, G. E., Meneses, R. M. & Rincón, D. A. (2012). La Investigación Evaluativa y su perspectiva en la Acreditación y Evaluación de Programas e Instituciones en Educación Superior, *XIII Asamblea General de la Asociación Latinoamericana de Facultades y Escuelas de Contaduría y Administración (ALAFEC)* (pp. 1-26). Buenos Aires, Argentina.
- Muñoz, A., Perez Zabaleta, A., Muñoz, A. & Sanchez, C. (2013). La evaluación de políticas públicas: una creciente necesidad en la unión europea. *Revista de Evaluación de Programas y Políticas Públicas*, 1, 1-30.
- Neuman, A., Shahor, N., Shina, I., Sarid, A. & Saar, Z. (2013). Evaluation utilization research -Developing a theory and putting it to use. *Evaluation and Program Planning*, 36, 64-70. DOI: <http://dx.doi.org/10.1016/j.evalprogplan.2012.06.001>
- Nicoletti, J. A. (2013). La evaluación de la calidad educativa. Investigación de base evaluativa en centros de educación superior. *Revista Argentina de Educación Superior*, 6, 189-202.
- Nitsch, M., Waldherr, K., Denk, E., Griebler, U., Marent, B. & Forster, R. (2013). Participation by different stakeholders in participatory evaluation of health promotion: A literature review. *Evaluation and Program Planning*, 40 (1), 42-54. DOI: <http://dx.doi.org/10.1016/j.evalprogplan.2013.04.006>
- Paricio, J. (2015). *Análisis de los modelos de calidad de la educación superior. Diseño de una metodología de análisis multidimensional*. Tesis Doctoral, Universidad de Zaragoza.
- Patel, M. (2002a). A meta-evaluation, or quality assessment, of the evaluations in this issue, based on the African Evaluation Guidelines: 2002. *Evaluation and Program Planning*, 25, 329-332.
- Patel, M. (2002b). The African Evaluation Guidelines: 2002. A checklist to assist in planning evaluations, negotiating clear contracts, reviewing progress and ensuring adequate completion of an evaluation. *Evaluation and Program Planning*, 25, 481-492.
- Patton, M. Q. (2012). A utilization-focused approach to contribution analysis. *Evaluation*, 18, 364-377.
- Perassi, Z. (2009). Commentary: Evaluar un Programa Educativo: Una Experiencia Formativa Compleja, *Revista Iberoamericana de Evaluación Educativa*, 2(2), 172-195.
- Perez Juste, R. (2002). La evaluación de programas en el marco de la educación de calidad. *XXI Revista de Educación*, 4, 43-76.
- Perrin, B. (2001). Commentary: Making yourself -and evaluation- useful. *American Journal of Evaluation*, 22 (2), 252-259.
- Pinkerton, S. D., Johnson-Massoti, A. P., Derse, A. & Layde, P. M. (2002). Ethical issues in cost-effectiveness analysis. *Evaluation and Program Planning*, 25, 71-83.
- Preskill, H., & Boyle, S. (2008). A multidisciplinary model of evaluation capacity building. *American Journal of Evaluation*, 29, 443-459.
- Renger, R & Hurley, C. (2006). From theory to practice: Lessons learned in the application of the ATM approach to developing logic models. *Evaluation and Program Planning*, 29, 106-119.
- Rodriguez-Campos, L. (2012). Stakeholder involvement in evaluation: Three decades of the American Journal of Evaluation. *Journal of Multidisciplinary Evaluation*, 17, 57- 79.
- Rodríguez-Campos, L. (2012). Advances in collaborative evaluation, *Evaluation and Program Planning*, 35, 523-528. DOI: <http://dx.doi.org/10.1016/j.evalprogplan.2011.12.006>

- Roseland, D., Lawrenz, F. & Thao, M. (2015). The relationship between involvement in and use of evaluation in multi-site evaluations. *Evaluation and Program Planning*, 48, 75-82. DOI: <http://dx.doi.org/10.1016/j.evalprogplan.2014.10.003>
- Ryan, K. E. (2004). Serving public interests in educational accountability: Alternative approaches to democratic evaluation. *American Journal of Evaluation*, 25 (4), 443-460.
- Scheerens, J. (2004). The evaluation culture. *Studies in Educational Evaluation*, 30(2), 105-124.
- Schwandt, T.A. (2002). *Evaluation Practice Reconsidered*. New York NY. Peter Lang Publishing
- Schwartz, R. & Mayne, J. (2005). Assuring the quality of evaluative information: theory and practice. *Evaluation and Program Planning*, 28(1), 1-14.
- Schweigert, F., J. (2007). The priority of justice: A framework approach to ethics in program evaluation. *Evaluation and Program Planning*, 30, 394-399.
- Scriven, M. (2000). *The logic and methodology of checklists*. Recuperado de www.wmich.edu/evalctr/checklists/
- Scriven, M. (2003). Evaluation in the new millennium: The transdisciplinary vision. En S. I. Donaldson & M. Scriven (Eds.), *Evaluating Social Programs and Problems. Visions for the New Millennium*, (pp. 19-42). Mahwah, N. J.: Lawrence Earlbaum Associates
- Smith, M. J. (2010). *Handbook of Program Evaluation for Social Work and Health Professionals*. New York: Oxford University Press
- Sondergeld, T. & Koskey, K. (2011). Evaluating the impact of an urban comprehensive school reform: An illustration of the need for mixed methods. *Studies in Educational Evaluation*, 37, 94-107. DOI: <http://dx.doi.org/10.1016/j.stueduc.2011.08.001>
- Stake, R. (2006). *Evaluación comprensiva y evaluación basada en estándares*, Editorial Graó, Barcelona.
- Stufflebeam, D. L. (2000). *Guidelines for developing evaluation checklists*. Recuperado de www.wmich.edu/evalctr/checklists/
- Stufflebeam, D. L. (2001a). Interdisciplinary PhD Programming in Evaluation. *American Journal of Evaluation*, 22(3), 445-455.
- Stufflebeam, D. L. (2001b). The metaevaluation imperative. *American Journal of Evaluation*, 22(2), 183-209.
- Stufflebeam, D. L. (2004). A note on the purposes, development, and applicability of the Joint Committee Evaluation Standards. *American Journal of Evaluation*, 25, 1, 99-102.
- Taut, S. (2008). What have we learned about stakeholders involvement in program evaluation. *Studies in Educational Evaluation*, 34, 224-230.
- Thomas, V. G. & Madison, A. (2010). Integration of Social Justice Into the Teaching of Evaluation. *American Journal of Evaluation*, 31 (4), 570-583.
- Urban, J. B., Hargraves, M. & Trochim, W. M. (2014). Evolutionary Evaluation: Implications for evaluators, researchers, practitioners, funders and the evidence-based program mandate. *Evaluation and Program Planning*, 45, 127-139. DOI: <http://dx.doi.org/10.1016/j.evalprogplan.2014.03.011>
- Urban, J. B. & Trochim, W. M. (2009). The role of evaluation in research-practice integration: Working toward the "Golden spike". *American Journal of Evaluation*, 30 (4), 538-553. DOI: <http://dx.doi.org/10.1177/1098214009348327>
- Vanhoof, J. & Van Petegem, P. (2010). Evaluating the quality of self-evaluations: The (mis)match between internal and external meta-evaluation. *Studies in*

- Educational Evaluation*, 36, 20–26. DOI: <http://dx.doi.org/10.1016/j.stueduc.2010.10.001>
- Walton, M. (2014). Applying complexity theory: A review to inform evaluation design. *Evaluation and Program Planning*, 45, 119-126. DOI: <http://dx.doi.org/10.1016/j.evalprogplan.2014.04.002>
- Wasserman, D. L. (2010). Using a systems orientation and foundational theory to enhance theory-driven human service program evaluations. *Evaluation and Program Planning*, 33, 67-80. DOI: <http://dx.doi.org/10.1016/j.evalprogplan.2009.06.005>
- Weiss, C. H. (2004). On theory-based evaluation: Winning friends and influencing people. *The Evaluation Exchange*, 9(4), 1-5.
- White, H. (2013). The Use of Mixed Methods in Randomized Control Trials. *New Directions for Evaluation*, 138(2), 61-73. DOI: <http://dx.doi.org/10.1002/ev.20058>
- Youker, B. W., Ingraham, A. & Bayer, N. (2014). An assessment of goal-free evaluation: Case studies of four goal-free evaluations. *Evaluation and Program Planning*, 46, 10-16. DOI: <http://dx.doi.org/10.1016/j.evalprogplan.2014.05.002>
- Yusa, A., Hynie, M. & Mitchell, S. (2016). Utilization of internal evaluation results by community mental health organizations: Credibility in different forms. *Evaluation and Program Planning*, 54, 11-18. DOI: <http://dx.doi.org/10.1016/j.evalprogplan.2015.09.006>

Autor

To know more / Saber más

Escudero, Tomás (tescuder@unizar.es).

Catedrático de Universidad, Área de Métodos de Investigación y Diagnóstico en Educación, Departamento de Educación, Instituto de Ciencias de la Educación (Universidad de Zaragoza). Dirección postal: Universidad de Zaragoza. Instituto de Ciencias de la Educación. C/ Pedro Cerbuna, 12. 50009 Zaragoza (España)



Revista Electrónica de **I**nvestigación y **E**valuación **E**ducativa
E-Journal of Educational Research, Assessment and Evaluation

[ISSN: 1134-4032]

© Copyright, RELIEVE. Reproduction and distribution of this articles it is authorized if the content is no modified and their origin is indicated (RELIEVE Journal, volume, number and electronic address of the document).

© Copyright, RELIEVE. Se autoriza la reproducción y distribución de este artículo siempre que no se modifique el contenido y se indique su origen (RELIEVE, volumen, número y dirección electrónica del documento).